

Atomic Structure of *Mycobacterium tuberculosis* CYP121 to 1.06 Å Reveals Novel Features of Cytochrome P450*

Received for publication, September 27, 2002, and in revised form, October 31, 2002
Published, JBC Papers in Press, November 14, 2002, DOI 10.1074/jbc.M209928200

David Leys†§¶, Christopher G. Mowat‡**, Kirsty J. McLean‡§, Alison Richmond‡‡, Stephen K. Chapman**, Malcolm D. Walkinshaw‡, and Andrew W. Munro‡§

From the ‡Department of Biochemistry, University of Leicester, The Adrian Building, University Road, Leicester LE1 7RH, United Kingdom, the ‡Structural Biochemistry Group, Institute of Cellular and Molecular Biology, Michael Swann Building, University of Edinburgh, The King's Buildings, West Mains Road, Edinburgh EH9 3JR, United Kingdom, the ‡‡Department of Pure and Applied Chemistry, University of Strathclyde, The Royal College, 204 George Street, Glasgow G1 1XL, United Kingdom, and the **Department of Chemistry, University of Edinburgh, The King's Buildings, West Mains Road, Edinburgh EH9 3JJ, United Kingdom

The first structure of a P450 to an atomic resolution of 1.06 Å has been solved for CYP121 from *Mycobacterium tuberculosis*. A comparison with P450 EryF (CYP107A1) reveals a remarkable overall similarity in fold with major differences residing in active site structural elements. The high resolution obtained allows visualization of several unusual aspects. The heme cofactor is bound in two distinct conformations while being notably kinked in one pyrrole group due to close interaction with the proline residue (Pro³⁴⁶) immediately following the heme iron-ligating cysteine (Cys³⁴⁵). The active site is remarkably rigid in comparison with the remainder of the structure, notwithstanding the large cavity volume of 1350 Å³. The region immediately surrounding the distal water ligand is remarkable in several aspects. Unlike other bacterial P450s, the I helix shows no deformation, similar to mammalian CYP2C5. In addition, the positively charged Arg³⁸⁶ is located immediately above the heme plane, dominating the local structure. Putative proton relay pathways from protein surface to heme (converging at Ser²⁷⁹) are identified. Most interestingly, the electron density indicates weak binding of a dioxygen molecule to the P450. This structure provides a basis for rational design of putative antimycobacterial agents.

The cytochromes P450 (P450s)¹ are a superfamily of heme-containing mono-oxygenases (1). They are famed for their roles in drug metabolism/detoxification and steroid synthesis in mammals (e.g. Refs. 2 and 3) but are also found in virtually all other life forms. The vast majority of P450s bind and cleave molecular oxygen in a two-step reaction involving delivery of

two electrons from a redox partner. In eukaryotic P450s, this is typically the diflavin enzyme NADPH cytochrome P450 reductase (4). However, in prokaryotes the reducing equivalents are usually supplied via two enzymes: NAD(P)H flavodoxin/ferredoxin reductase and either a ferredoxin or a flavodoxin mediator of electrons to the P450 (5, 6). Many bacterial P450s are of particular biotechnological and environmental relevance, given their role in degradation of recalcitrant organic molecules for use as energy sources (e.g. Refs. 7 and 8). Bacterial P450s have also proven to be experimentally tractable systems for elucidation of P450 structure and mechanism, due largely to the fact that they are soluble enzymes, as opposed to their eukaryotic membrane-associated counterparts (5). A breakthrough was made recently with the atomic structure of the first mammalian P450 (a form of P450 2C5). In this case, extensive protein engineering was required to remove the N-terminal membrane anchor peptide, to promote release from the cellular membrane and to prevent aggregation of the resultant solubilized P450 (9).

Whereas the genome sequence of *E. coli* is devoid of P450-encoding genes (10), the genomes of the industrially important bacterium *Streptomyces coelicolor* and the pathogen *Mycobacterium tuberculosis* (Mtb) indicate that there are 18 and 20 P450s encoded, respectively (11, 12). The genome sequence of Mtb revealed several interesting features, most notably the preponderance of genes involved in lipid metabolism; ~250 enzymes involved in lipid metabolism have been identified in Mtb (cf. ~50 for the similarly sized *Escherichia coli* genome). The importance of lipid metabolism in Mtb may explain the high P450 content, since typical cellular functions for P450s in eukaryotes and other bacteria are as mono-oxygenases for lipophilic fatty acids, polyketides, steroids, and xenobiotics (e.g. see Ref. 2). Sterols are rare in bacteria, but the P450 product of the Mtb Rv0764c gene (CYP51) is highly related to eukaryotic sterol demethylases and has been shown to catalyze demethylation of plant and eukaryotic sterols (13). The crystal structure of the Mtb CYP51 was solved in complex with the azole antifungal drug fluconazole (14). This structure is of biotechnological importance, given the importance of the eukaryotic CYP51s in physiology and also in view of the fact that the yeast and fungal CYP51 enzymes (lanosterol demethylases) are validated targets for the azole antifungal class of drugs of which fluconazole is a member (15). In recent studies, the potency of these drugs as antibiotics against growth of Mtb's relative *Mycobacterium smegmatis* and the related actinomycete *S. coelicolor* has been shown (16).

The biochemical functions of the remaining P450 enzymes in Mtb are not immediately obvious. This is primarily because of

* This work was supported by the Biotechnology and Biological Sciences Research Council, Engineering and Physical Sciences Research Council, and European Union (Project "X-TB"). Access at EMBL Hamburg was supported by the European Community-Access to Research Infrastructure Action of the Improving Human Potential Program to the EMBL Hamburg Outstation (Contract HPRI-CT-1999-00017). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The atomic coordinates and structure factors (code 1N40 (RCSB017490) and 1N4G (RCSB017506)) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (<http://www.rcsb.org/>).

§ Member of the TB Structural Genomics Consortium.

¶ To whom correspondence should be addressed. Tel.: 44-116-252-3484; Fax: 44-116-252-3473; E-mail: dl37@le.ac.uk.

¹ The abbreviation used is: P450, cytochrome P450.

TABLE I
 Data collection and phasing statistics for CYP121

Data set	Native	Iodopyrazole	PtCN	K ₂ PtCl ₄	Hg(NO ₃) ₂
Source	ESRF-ID14.4	DESY-X11	SRS 14.1	Rotating anode	Rotating anode
Resolution (Å)	10–1.06	20–1.8	20–2.5	20–4.3	20–2.8
Unique observations	185,472	44,862	16,817	3544	11,923
Completeness (%)	88.5	95.9	97.2	98.1	95.5
R _{merge} (%) ^a	9.2	7.2	5.1	8.4	7.8
I/I	6.5	8.1	28.9	9.2	12.7
Temperature (K)	107	107	107	107	107
Heavy atom sites		2	3	3	1
FOM 2.0–2.3 Å	0.42				

^a $R_{\text{merge}}(I) = \sum_h \sum_i |I_{h,i} - \langle I_h \rangle| / \sum_h \sum_i I_{h,i}$ where I is the observed intensity. $\langle I \rangle$ is the average intensity of multiple observations of symmetry-related reflections.

limited similarity to other members of the P450 superfamily with known catalytic properties. Most show highest similarity to other P450 enzymes in Mtb, perhaps suggesting co-evolution as oxygenases of related, complex lipids in the bacterium (17). We decided to focus efforts on the product of gene *Rv2276*, a P450 enzyme classified as CYP121 (1). This P450 shows amino acid sequence similarity to a number of polyketide mono-oxygenase P450s, including the structurally characterized P450 EryF (CYP107A1) from *Saccharopolyspora erythraea* (18), suggesting a potential role in polyketide metabolism in Mtb. In addition, the known high affinity of P450 EryF for the azole antifungal class of drugs suggests that CYP121 could be a potential drug target in the pathogen, particularly since the role of CYP51, its validity as a drug target, and whether sterol metabolism is really a feature of the pathogen remain to be established. Indeed, the fact that a Δ CYP51 strain of *S. coelicolor* is nonlethal (16, 19) and that the strain exhibits no significant alteration in sensitivity to azole antifungal drugs (16) confirms that CYP51 is not the target in this actinomycete.

Our initial expression and biophysical studies have established that the Mtb P450 CYP121 has properties typical for members of this class (20). Whereas the physiological role remains unclear, the enzyme binds bulky azole antifungal drugs with high affinity, and the binding constants for these drugs are perturbed by the presence of erythromycin and other large polyketides, suggesting that CYP121 may metabolizes polyketides or bulky polycyclics *in vivo*. However, what is most remarkable is that in comparison with Mtb CYP51, the binding of a variety of azole antifungal drugs to Mtb CYP121 is tighter and that the order of their K_d values correlates well with the minimal inhibitory concentration values for these drugs for *M. smegmatis* and *S. coelicolor* (16). Since the drugs retain their high potency against a *S. coelicolor* CYP51 knockout strain, other P450 enzyme(s) must be the true targets in this bacterium; thus, the fact that potency of azole drugs *versus* mycobacteria correlates with CYP121 K_d values must make this a prime drug target candidate (20). We have previously reported crystallization of CYP121 (21), and in this paper we report the crystal structure of CYP121 to 1.06 Å. This unprecedented level of resolution for a P450 system allows, for the first time, a truly atomic description of the oxygen scission site.

EXPERIMENTAL PROCEDURES

Expression and Purification of CYP121—The *Rv2276* gene encoding CYP121 was cloned by PCR from a Mtb chromosomal DNA library (cosmid MTCY339, obtained from Prof. Stewart Cole at the Institut Pasteur, Paris, France) and cloned into expression plasmid pET11a to produce clone PKM2b, as described previously (20). The CYP121 protein was produced in *E. coli* strain HMS174 (DE3) with growth of the culture after isopropyl-1-thio- β -D-galactopyranoside induction performed at low temperature (18 °C) to promote production of soluble enzyme. The cells were broken using a French press and sonication, and CYP121 was purified to homogeneity by ammonium sulfate fractionation (30–70% fraction retained), followed by successive column chro-

matography steps on phenyl-Sepharose, Q-Sepharose, and hydroxyapatite resins, as described (20).

Crystallization and Structural Elucidation of CYP121—CYP121 crystals belonging to space group P6₃22 (unit cell dimensions $a = 77.81$ Å, $c = 263.82$ Å) were obtained as described previously (21). A complete native data set up to 1.06 Å could be measured on a single flash-cooled crystal on ID14–4 at the European Synchrotron Radiation Facility. Different sections of the crystal were exposed to avoid excessive radiation damage due to the intense beam. Whereas these crystals clearly diffract further, both the long cell axis of 263 Å and the rapid decay of diffraction quality hampered the collection of higher resolution reflections. Data were processed and scaled with the HKL package programs DENZO and SCALEPACK (22). Despite the availability of several P450 structures, the crystal structure could not be solved using molecular replacement and was solved using a Multiple Isomorphous Replacement with Anomalous Scattering (MIRAS) approach instead. Data were collected on station 14–1 at Synchrotron Radiation Source for two independent heavy metal derivatives. Initial heavy atom sites were found using the program RSPS (23) and used for detection of minor sites by difference Fourier analysis. Heavy atom sites were refined using the program MLPHARE (24), and the weak anomalous signal of the iron was included in the final phase calculations. Data collection and phasing statistics are given in Table I. The program ARP-WARP (25) was used to build an initial model in the solvent flattened maps. This initial model was manually adjusted using TURBO-FRODO (26) and refined using REFMAC5 (27). The structure of the uncomplexed enzyme was refined with anisotropic B -factors for all atoms including a riding hydrogen atom model. With successive rounds of refinement and model building, the entire CYP121 sequence was fitted to the electron density map with the exception of the N-terminal three residues. Multiple conformations were observed for 33 of the 395 residues defined. The final R -factor was 0.132 (R -free = 0.153) for the model consisting of the protein, 800 water molecules, and 3 sulfates per asymmetric unit. The enzyme was complexed with hydrophilic compounds by soaking in the mother liquor (2.3 M ammonium sulfate, pH 6.5) supplemented with a large excess of the ligand. In order to obtain data on complexes with hydrophobic compounds, the crystals were placed in 30% polyethylene glycol 8000, 0.2 M ammonium sulfate, pH 6.5, supplemented with the compound of interest. Only the hydrophilic iodopyrazole proved to bind in the active site, and data were collected for complexed crystals to 1.8 Å. Final refinement statistics for the two crystal structures are given in Table II.

RESULTS AND DISCUSSION

General Structural Features—Globally, the CYP121 structure resembles a triangular prism, as observed for all other P450 structures to date. As with other P450s, there are two distinct structural “domains,” with the heme sandwiched between a smaller β sheet-rich module and a larger helix-rich domain (Fig. 1). The P450s have relatively few definitive amino acid motifs, and evolutionarily distinct P450s may exhibit as little as 15% amino acid identity. However, the cysteine ligand to the heme iron and the region immediately surrounding this amino acid (the “heme-binding” motif) at the C-terminal section of the P450 are strongly conserved (Fig. 2).

Overall Enzyme Fold—Despite the elucidation of several P450 crystal structures, new cytochrome P450 models often reveal distinct structural features that are unpredictable from primary sequence alignments. Given the relative low homology

TABLE II
Final refinement statistics for the CYP121 crystal structures

Data set	Native	Iodopyrazole
Space group	P6 ₅ 22	P6 ₅ 22
Resolution (Å)	10–1.06	20–1.8
R_{cryst}^a (%)	13.2	20.5
R_{free}^a (%)	15.3	21.7
Number of non-hydrogen atoms	3931	3536
Average B -factor	10.34	15.2
r.m.s. ^b deviations in bond lengths (Å)	0.018	0.024
r.m.s. deviations in bond angles (degrees)	1.94	2.35

^a $R_{\text{cryst}} = \sum ||F_o| - |F_c|| / \sum |F_o|$; R_{free} is the same as R_{cryst} but was calculated using a separate validation set of reflections that was excluded from the refinement process.

^b r.m.s., root mean square.

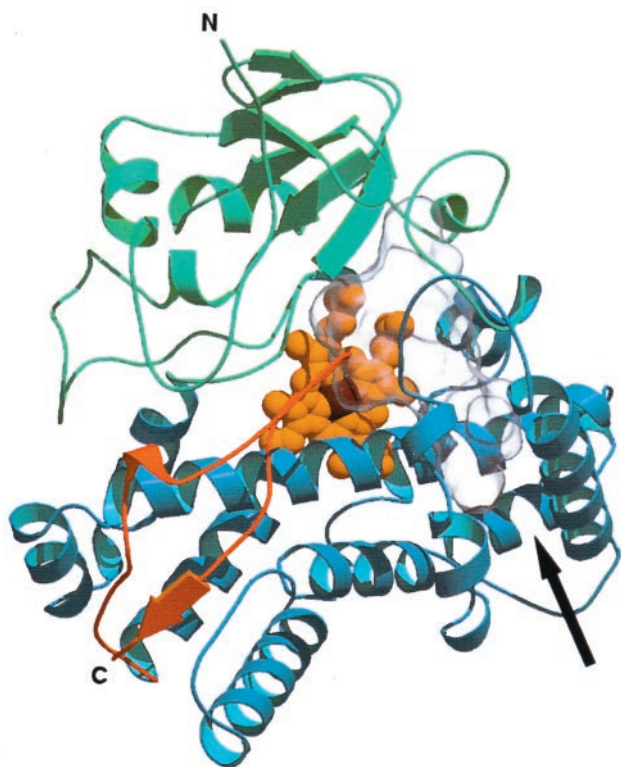


FIG. 1. Overall fold of *Mtb* CYP121. The atomic structure of CYP121 is composed mainly of α helices. The heme (in space fill) is sandwiched between the two major domains of the structure: an α helix-rich domain (dark gray) and the smaller β sheet-rich domain (light gray). The large water-filled active site cavity of the enzyme is shown as a transparent surface. An arrow indicates the possible site of substrate entry between helices F and G.

(25.3% identity over 396 amino acids) with P450 EryF (CYP107A1; a macrolide monooxygenase) and the unsuccessful molecular replacement, the high three-dimensional similarity of the substrate-free CYP121 with that of substrate/inhibitor-bound CYP107A1 comes as a surprise (28, 29). Notwithstanding this high similarity in fold (Z score 42.3, root mean square deviation 2.6 Å for 379 C- α atoms), several distinct features can be observed in CYP121 for polypeptide stretches lining the active site cavity.

Upon comparison of CYP121 with CYP107A1 from N to C terminus, the first region that is drastically different is that immediately following the conserved B helix (Fig. 3). Starting at residue Met⁶², the polypeptide chain proceeds first with a 3_{10} helix and then with an extended loop forming the dome of the

CYP121	Mtu	Y M A A R R Q Q P F C S A A G R H A	355
CYP7A1	Human	Y M F P G G A T I C R L A A I H I	454
CYP7A1	Rabbit	Y M F P G G A T I C R L A A I H I	478
CYP128	Mtu	F M A P G G C I L L C T Q A A M E G	445
CYP2D6	Human	F M P S A A R R A C P F A A R H E L	453
CYP2C5	Human	F M P S A A R R M C P F A A R H E L	442
CYP107A1	Ser	F M P S A A R R I C P F A A R H E L	361
CYP107H1	Bsu	F M P S A A R R I C P F A A R H E L	355
CYP119	Sao	F M P S A A R R I C P F A A R H E L	327
CYP130	Mtu	I L T S H A A H L C A A A A R H C	364
CYP104	Atu	Q S S G G C P R H C A Q I S R Q T V	366
CYP3A4	Human	Y T P S G G P R N C M R A N L N M	451
CYP102A1	Bme	F K P P G G R R I C Q Q S A A L H E A	410
CYP132	Mtb	Y P S G G C R R I C Q S A A L N M	419
CYP51	Mtb	W I P S A A R R C P A A A A I Q I	404

FIG. 2. The heme binding motif in *M. tuberculosis* CYP121 and other cytochromes P450. The heme-binding motif region for *Mtb* CYP121 (10 amino acids either side of the cysteine ligand to the heme iron) is shown aligned versus those from prokaryotic and eukaryotic P450s using the ClustalW program (EBI). Four amino acids are conserved throughout, including the cysteine (Cys³⁴⁵ in CYP121), and are highlighted in white against a black background. These identical residues are also indicated by an asterisk at the base of the alignment, and other positions where the amino acid class is retained in each P450 are indicated by dots. Highlighted in gray are those individual amino acids present in 40% or more of the P450s at the respective positions. P450 enzymes are indicated by their formal classification “CYP” numbers (1) and by their species of origin. *Mtu*, *M. tuberculosis*; *Ser*, *S. erythraea*; *Bsu*, *B. subtilis*; *Sao*, *S. solfataricus*; *Atu*, *Agrobacterium tumefaciens*; *Bme*, *B. megaterium*.

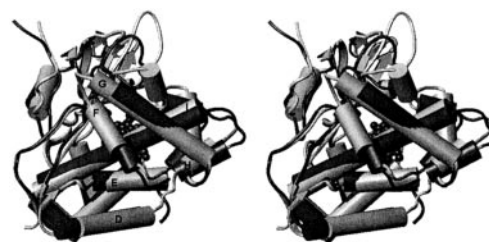


FIG. 3. Structural comparison of CYP121 with CYP107A1 (P450 EryF) from *S. erythraea*. Shown is a stereo view of the overlay of *Mtb* CYP121 (dark gray) with *S. erythraea* CYP107A1 (light gray). Loops and secondary elements lining the active site cavity are clearly more divergent than the remainder of the structure.

active site relative to the heme plane. This loop proceeds into a second 3_{10} helix that is immediately followed by helices B' and C. The corresponding polypeptide region in CYP107A1 is longer and bears little or no resemblance, having no regular 3_{10} helical structure, whereas helix B' is separated from C by several residues. The C helix in CYP121 is immediately adjacent to helix B' and makes a different angle with other conserved structural elements. It is separated from the short D helix by a flexible region that precedes Glu¹¹⁰. From this position onward, both polypeptide chains again assume very similar positions. After helix D and E, the helices F and G are slightly reoriented with respect to the remainder of the structure, and the FG loop that forms a part of the active site dome is bent inward with respect to the CYP107A1 structure. Whereas the CYP121 helix H still occupies a distinct conformation, as does the beginning of helix I, the polypeptide chain trace is once again highly similar to the main part of helix I and remains so for the rest of the structure. Two minor exceptions are the loop following the K helix and the C-terminal region, which are both slightly reoriented and form part of the dome of the active site. These reorganizations in the secondary elements and loops lining the active site cavity result in a slightly increased volume for the CYP121 active site of 1337 Å compared with 1115 Å for the substrate-bound CYP107A1. Whereas in CYP121 the main part of the active site is more compressed due to the closer packing of loops forming the dome of the active site, a significant part of the cavity extends between helices F and G (Fig. 1), accounting for the increased volume in CYP121. As has been observed for the majority of P450 crystal structures (with the notable exceptions of *Mtb* CYP51 and *Bacillus megaterium* P450 BM3) (14, 30), the active

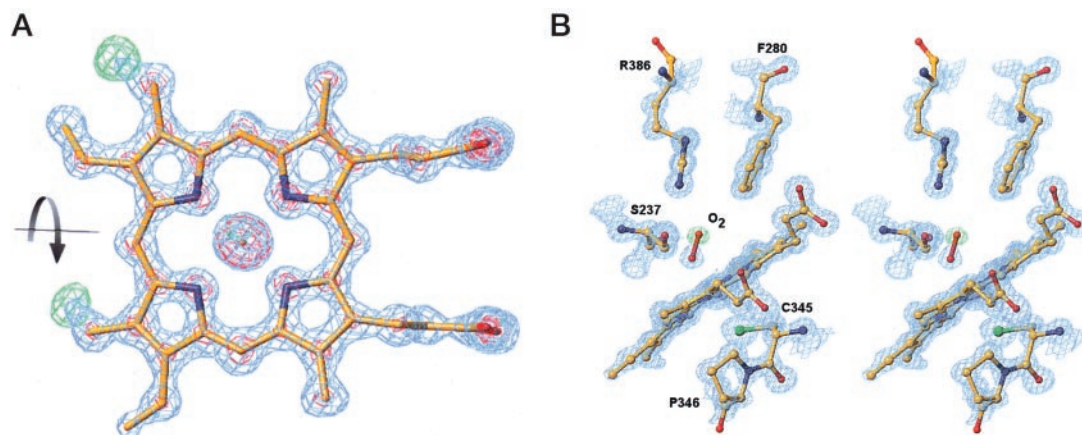


FIG. 4. *A*, different heme conformations in CYP121. The main conformation of the heme *B* group of CYP121 is shown, surrounded with 1.5 (blue) and 4 (purple) σ contoured 2FoFc electron density. The additional FoFc density (contoured at 3 σ in green) and the lack of strong 2FoFc density of the vinyl atoms of the main conformation clearly indicate a second conformer. The two different conformers are related by the indicated 180° rotation about the heme plane. *B*, stereo view of the active site of CYP121. Residues surrounding the heme are shown in addition to the putative oxygen ligand. The corresponding 2FoFc electron density is contoured at 1.5 σ in blue. The heme is notably kinked due to the close interaction with Pro³⁴⁶. The FoFc density is contoured at 3 σ in green and reveals a second atom close to the position of the sixth water ligand.

site is apparently inaccessible from the surface in this substrate-free form of the enzyme. This obviously raises the question of how the putative substrate reaches the active site. As has been proposed for other P450s, we envisage the CYP121 structure as oscillating between open and closed conformations whereby the “breathing motions” of the protein, specifically movements of the secondary elements and loops lining the active site, allow for transient exposure of gaps allowing substrate access. It is especially interesting that the active site seems to protrude in between helices F and G, allowing small movements in the relative positions of these helices to open an active site channel (Fig. 1). Similar motions in helices F/G and the interconnecting loop region have been demonstrated to occur for the CYP119 monooxygenase from *Sulfolobus solfataricus* (31).

P450 Heme Binding in Atomic Detail—As in other P450 enzymes, a cysteine (Cys³⁴⁵) provides the proximal axial ligand to the heme iron, with a well defined water molecule as the distal axial ligand. The Fe-S distance is 2.30 Å, whereas the Fe-O distance is 2.21 Å, both values being highly similar to those in other P450 structures obtained.

An unusual aspect of the CYP121 structure is the fact that the heme is clearly bound in one of two distinct conformations, related by a 180° rotation through an axis of symmetry across the $\text{CH}_\alpha\text{-Fe-CH}_\delta$ atoms of the molecule. In short, the heme is bound either in the “a” or “b” conformer, a phenomenon well known for other heme proteins (e.g. cytochrome *b*₅) (e.g. see Ref. 32). This mixture of conformers does not otherwise impact on the overall fold of the enzyme. One of the conformers predominates, and the relative occupancy is refined to a 70:30 mixture of the species (Fig. 4*a*). In cytochrome *b*₅, there is dynamic exchange of the heme, such that the proportions of the a- and b-type conformers can be perturbed considerably by temperature (e.g. see Ref. 33). The situation is extremely unlikely to be the same in the cytochromes P450, where the heme is bound deep in the protein (not superficially as in *b*₅) and requires protein denaturation to be released. Whereas insertion of heme and modified hemes into P450 has been demonstrated, this was achieved only very slowly by denaturation of the P450 and refolding by gradual removal of denaturant after the addition of the exogenous heme (e.g. see Ref. 34). The situation is highly unlikely to occur *in vivo*, and we infer that the heme-binding step in CYP121 (and other P450s) occurs at a relatively early stage in the protein folding process, and the position of the heme as bound commits the P450 to this conformation throughout its lifetime *in vivo*.

One of the four pyrrole rings of the heme is distorted out-of-plane by the close interaction with the side chain of Pro³⁴⁶. The pyrrole is kinked toward the distal face by an angle of ~30° (Fig. 4*b*). Pro³⁴⁶ immediately follows the cysteinyl heme ligand (Cys³⁴⁵), and this Cys-Pro motif is found in a number of other cytochrome P450 enzymes, with several of the *Mtb* P450s having this feature (Fig. 2). This structural perturbation to heme planarity may therefore well be a feature common in other P450 enzymes but previously unrecognized due to the absence of the Cys-Pro motif in P450 structures solved to date. This structural element could indeed be essential to the specific electronic properties of the heme in these oxidases. Recently, we have demonstrated in the P450 BM3 system that a phylogenetically conserved Phe residue (Phe³⁹³ located 7 amino acids before the cysteine ligand) that also stacks with both heme plane and packs with the cysteinyl ligand exerts a large effect on the heme electronic properties (Fig. 2) (35). It is therefore highly likely that this proline residue, being closer to both heme and cysteinyl ligand than even the phenylalanine (Phe³³⁸ in CYP121), has a considerable impact on the heme catalytic and thermodynamic properties.

The Oxygen Scission Site to Atomic Detail—The water ligand to the heme in the resting state of the enzyme is hydrogen bonding to Ser²³⁷. This residue replaces the more general threonine residue found in cytochromes P450, which has been implicated in oxygen binding and/or proton delivery (36, 37). In turn, Ser²³⁷ hydrogen-bonds to the carbonyl backbone of Ala²³³ and to the side chain of Arg³⁸⁶ (Figs. 4*b* and 5). Whereas a distinctive deviation of conventional α -helix hydrogen bonding patterns in the region of the water ligand-contacting residues can be seen in all P450s of known structure, with the exception of CYP2C5 (9), the I helix in CYP121 is devoid of any significant alteration of hydrogen bonding patterns and resembles the eukaryotic CYP2C5 in this region. The presence of the bulky and positively charged residue Arg³⁸⁶ immediately adjacent to the oxygen scission site is unique to CYP121 and makes it one of the key residues in this enzyme (Fig. 4*b*). The side chain of Arg³⁸⁶ is wedged in between Ile²³⁶ and Phe²⁸⁰, whereas hydrogen bonding to Ser²³⁷, Gln³⁸⁵, and two water molecules in the active site cavity. The close proximity of the arginine and serine residues may define a binding site for a negatively charged functional group, possibly a carboxylate, as observed previously for e.g. P450 BM3 (CYP102A1) with its fatty acid substrates (38, 39).

In the last stages of refinement, additional density became

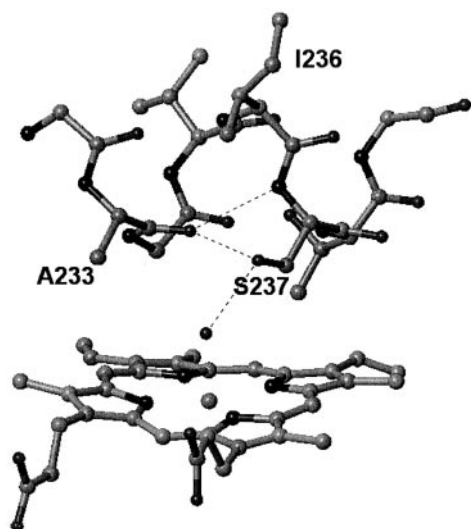


FIG. 5. **Hydrogen bonding pattern in the CYP121 I helix.** Residues of the CYP121 I helix close to the heme group are shown in *gray ball and stick* mode. All bacterial cytochromes P450 show a severe disruption of the hydrogen bonding pattern between Ala²³³ and Ser²³⁷ (CYP121 numbering). Remarkably, this distortion is not seen in CYP121 in the vicinity of the relevant amino acid (Ser²³⁷), similar to the mammalian CYP2C5 structure (9).

apparent immediately adjacent to the heme water ligand. The shape of this density is nearly spherical, and the center is located 1.2 Å from the sixth ligand (Fig. 4b). Only two models seem to fit this density equally well: either the sixth ligand occupies two distinct and mutually exclusive conformations, or a diatomic species with ~1.2-Å bond length (*e.g.* dioxygen) is bound to a proportion of the heme groups in the crystal, effectively superposing the densities of water ligation and dioxygen ligation. The presence of hydrogen peroxide can be excluded on the basis that the bond length of the observed species is significantly shorter than that of hydrogen peroxide (1.5 Å). The immediate environment of the additional density does not, however, correlate with the first model (*i.e.* two distinct water conformations), since there are no significantly stabilizing structural elements present to explain the fact that the ligating water molecule would preferentially bind outside, but close to, the ligation sphere of the iron. Therefore, we favor the second model (*i.e.* bound dioxygen), with an angle of nearly 120° between dioxygen and heme plane highly similar to other dioxygen-heme structures (*e.g.* see Ref. 40). However, it is unprecedented that dioxygen binds to ferric heme, and the possibility exists that the crystal might have partially reduced in the intense beam. However, during data collection, several parts of the crystal were exposed each for less than 10 s, minimizing the chance of significant reduction due to repetitively exposing fresh, fully oxidized parts of the crystal. Additionally, several data collections at medium resolution involving significantly longer exposures and higher irradiation doses in similar conditions did not reveal any additional density. Second, the putative bond length for the dioxygen-ferric ion is 2.2 Å, which is highly similar to the water-ferric ion distances but in significant deviation for dioxygen-ferrous ion distances of 1.8 Å. Therefore, we conclude that P450 enzymes might weakly bind dioxygen in the ferric form, a consequence of the electron rich character conferred upon the heme by the cysteinate ligation.

The putative dioxygen species forms an ideal model for analyzing the initial steps of the catalytic cycle, during which both electrons and protons need to be consecutively passed onto the ligand to break the dioxygen bond and create the highly reactive oxidizing intermediate. Upon initial reduction of the

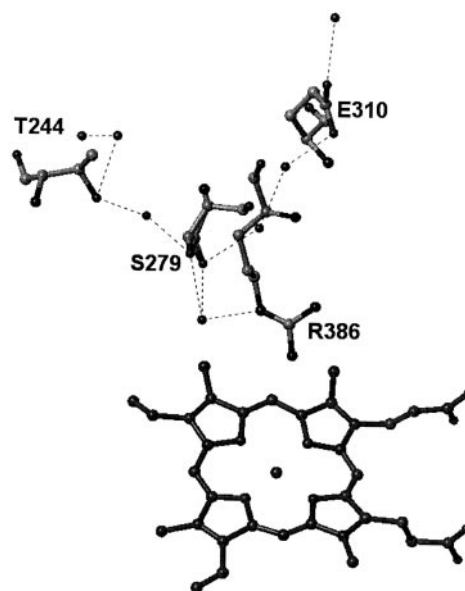


FIG. 6. **Putative proton relays in CYP121.** Shown is a view of two putative proton pathways leading from the solvent into the active site. Both pathways are mainly made by buried water molecules in addition to a single amino acid (Thr²⁴⁴ and Glu³¹⁰, respectively) in the two pathways. These pathways merge at Ser²⁷⁹ that occupies two distinct conformations, each contacting a water molecule that is hydrogen-bonded to Arg³⁸⁶. The latter residue is proposed to donate protons either directly or via Ser²³⁷ (not shown) to the oxygen intermediates.

heme, the dioxygen molecule can either accept a proton from proximal water molecules or from Ser²³⁷. However, we assume that the active site will be devoid of water molecules in the vicinity of the heme plane when complexed with substrate. Ser²³⁷, in turn, can easily receive a proton from Arg³⁸⁶. This residue contacts a water molecule buried in the interior of the protein. At this point, there are several other residues contacting this water molecule that could provide a proton. Most notable is Ser²⁷⁹, which occupies two distinct conformations in the CYP121 structure. Each of these conformations is in contact with a different series of residues and buried water molecules that eventually lead up to the surface of the molecule. Thus, it appears plausible that there may be a bifurcated proton delivery pathway from the P450s surface, with pathways converging at Ser²⁷⁹. This would have the potential to transport one proton at a time through each individual path (Fig. 6).

Structure of the Substrate Binding Site—The active site cavity of 1350 Å³ is filled with water molecules and is remarkably rigid. Whereas a significant proportion of both surface-exposed and interior residues exhibit multiple conformations, all of the residues involved in lining the active site adopt a single conformation. The dome of the active site chamber is 12 Å above the heme plane, being most open above pyrrole ring D (main conformation). The heme is 62 Å² solvent-exposed with access to pyrroles B and C limited by residues from the I helix. The majority of residues lining the active site are hydrophobic with the notable exceptions of Arg³⁸⁶ and Gln³⁸⁵. Whereas several antifungal agents bind tightly to CYP121 in solution (20), we have been unable to date to obtain crystallographic complexes with these compounds by either soaking native crystals or conducting co-crystallization trials. This is highly likely due to the extreme insolubility of the polycyclic azoles in the ammonium sulfate mother liquor but could also be due to restricted access of the azoles to the heme due to obstruction by Arg³⁸⁶. The single azole compound that was proven to bind into the active site cavity by soaking the crystals was iodopyrazole. This clearly demonstrates the fact that, even in the tight crystal

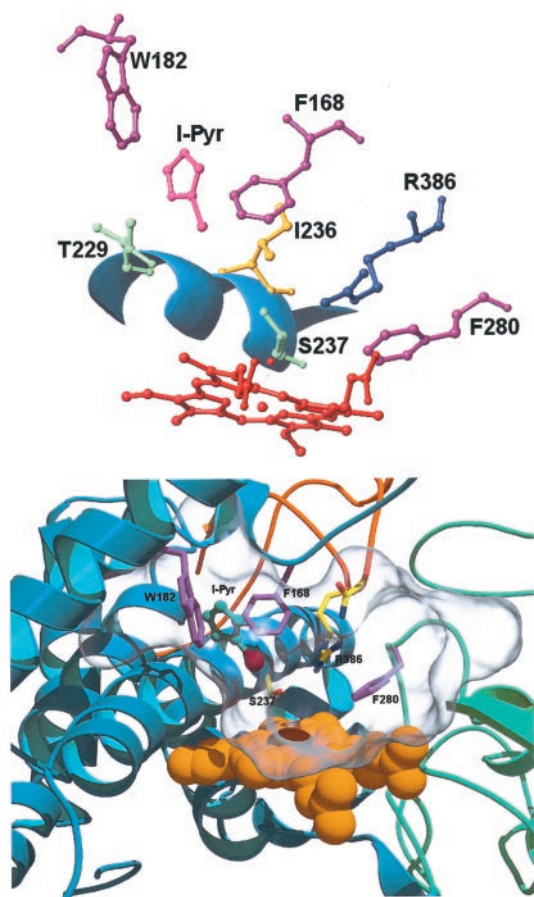


FIG. 7. **The iodopyrazole complex of CYP121.** A, key residues of the CYP121 active site in complex with iodopyrazole. Residues are depicted in ball and stick mode, color-coded according to property as follows. Blue, basic; purple, aromatic; yellow, hydrophobic; green, hydrophilic. The iodopyrazole (I-Pyr) ligand is shown in pink, and the heme group is shown in red. Part of the I helix is shown as a blue ribbon. B, overall view of the active site in complex with iodopyrazole. The active site cavity is depicted as a gray transparent surface with key residues shown as sticks. Secondary structure elements are color-coded according to the legend to Fig. 1.

packing, breathing motions must allow transient access from the outside into the otherwise inaccessible active site cavity. Surprisingly, the pyrazole compound does not ligate to the heme iron but binds instead in the small channel region between helices F and G. The pyrazole plane stacks at ~ 4 Å with Trp¹⁸² on one side and Phe¹⁶⁸ on the other side (Fig. 7, a and b). The bulky iodide atom points toward the active site cavity and is buried between the hydrophobic side chains of Phe¹⁶⁸, Thr²²⁹, and Ala²³³. A similar unexpected observation was made upon complexation with 4-phenylimidazole. In this case, the compound bound to the surface of the protein instead of acting as the sixth heme ligand. In spectral titrations, 4-phenyl imidazole was shown to bind to the heme iron, giving a typical (and complete) shift of the Soret maximum to ~ 423 nm. This clearly demonstrates that CYP121 can adopt conformations in solution distinct from that observed in the crystal and that these solution conformations are compatible with ligand binding to the iron.

CONCLUSIONS

CYP121 is similar in overall fold to other structurally characterized P450s. It contains unique structural elements involved in lining the active site. The heme binding and oxygen scission sites are particularly interesting, since these reveal several new features that could be general to P450s. First, the heme cofactor is bound in two distinct, mutually exclusive

conformations. Furthermore, it is contorted due to close interaction with the Cys³⁴⁵-Pro³⁴⁶ motif, a common structural element in Mtb P450s and present also in other prokaryotic and eukaryotic forms. A series of hydrogen-bonded amino acid and water molecules define two potential proton delivery pathways, and these apparently converge at Ser²⁷⁸, which has distinct conformations in the CYP121 structure. Most interestingly, the oxygen scission site is markedly different from other bacterial P450s, and, while resembling the eukaryotic CYP2C5 structure, it contains a unique arginine residue (Arg³⁸⁶). The atomic resolution obtained has allowed visualization of a species (probably dioxygen) weakly bound to the ferric heme iron.

Future studies on CYP121 will involve establishing its substrate selectivity through screening for activity against mycobacterial lipid extracts (and possibly infected tissue samples) and use of rational mutagenesis to define roles of key amino acids identified from this structural study in determining the biophysical and catalytic properties of the P450. A further priority is the solution of an azole-ligated CYP121 structure, since this may provide critical information required to facilitate *de novo* design of novel and more highly specific azole drugs that are potent anti-mycobacterials but that lack the cross-reactivity with human isoforms. Through such an approach, anti-P450 drugs may prove useful new agents in the war against multidrug-resistant strains of *M. tuberculosis*, which the World Health Organization describes as having the potential to cause a "global catastrophe."

Acknowledgments—We are grateful to Prof. Stewart Cole (Institut Pasteur, Paris) for provision of cosmid DNA and for helpful scientific discussions. We gratefully acknowledge beamtime received at Synchrotron Radiation Source (Daresbury, UK), European Synchrotron Radiation Facility (Grenoble, France), and EMBL X11 beamline at the DORIS storage ring, Deutsches Elektronen-Synchrotron (Hamburg, Germany).

REFERENCES

- Nelson, D. R., Koymans, L., Kamataki, T., Stegeman, J. J., Feyereisen, R., Waxman, D. J., Waterman, M. R., Gotoh, O., Coon, M. J., Estabrook, R. W., Gunsalus, I. C., and Nebert, D. W. (1996) *Pharmacogenetics* **6**, 1–42
- Guengerich, F. P. (2001) *Chem. Res. Toxicol.* **14**, 611–650
- Imai, T., Globberman, H., Gertner, J. M., Kagawa, N., and Waterman, M. R. (1993) *J. Biol. Chem.* **268**, 19681–19689
- Wang, M., Roberts, D. L., Paschke, R., Shea, T. M., Masters, B. S. S., and Kim, J. J. P. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94**, 8411–8416
- Munro, A. W., and Lindsay, J. G. (1996) *Mol. Microbiol.* **20**, 1115–1125
- Unno, M., Christian, J. F., Sjödin, T., Benson, D. E., Macdonald, I. D. G., Sligar, S. G., and Champion, P. M. (2002) *J. Biol. Chem.* **277**, 2547–2553
- Gunsalus, I. C., and Wagner, G. C. (1978) *Methods Enzymol.* **52**, 166–188
- Boddupalli, S. S., Hasemann, C. A., Ravichandran, K. G., Lu, J. Y., Goldsmith, E. J., Dieneshofer, J., and Peterson, J. A. (1992) *Proc. Natl. Acad. Sci. U. S. A.* **89**, 5567–5571
- Williams, P. A., Cosme, J., Sridhar, V., Johnson, E. F., and McRee, D. E. (2000) *Mol. Cell.* **5**, 121–131
- Miles, C. S., Ost, T. W. B., Noble, M. A., Munro, A. W., and Chapman, S. K. (2000) *Biochim. Biophys. Acta* **1543**, 383–407
- Cole, S. T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S. V., Eiglmeier, K., Gas, S., Barry, C. E., Tekai, F., Badcock, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R., Devlin, K., Feltwell, T., Gentles, S., Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Krogh, A., McLean, J., Moule, S., Murphy, L., Oliver, K., O'Neil, S., Rabinowitsch, E., Rajandream, M. A., Rogers, J., Rutter, S., Seeger, K., Skelton, J., Squares, R., Squares, S., Sulston, J. E., Taylor, K., Whitehead, S., and Barrell, B. G. (1998) *Nature* **393**, 537–544
- Bentley, S. D., Chater, K. F., Cerdeno-Tarraga, A. M., Challis, G. L., Thomson, N. R., James, K. D., Harris, D. E., Quail, M. A., Kieser, H., Harper, D., Bateman, A., Brown, S., Chandra, G., Chen, C. W., Collins, M., Cronin, A., Fraser, A., Goble, A., Hidalgo, J., Hornsby, T., Howarth, S., Huang, C. H., Kieser, T., Larke, L., Murphy, L., Oliver, K., O'Neil, S., Rabinowitsch, E., Rajandream, M. A., Rutherford, K., Rutter, S., Seeger, K., Saunders, D., Sharp, S., Squares, R., Squares, S., Taylor, K., Warren, T., Wietzorrek, A., Woodward, J., Barrell, B. G., Parkhill, J., and Hopwood, D. A. (2002) *Nature* **417**, 141–147
- Bellamine, A., Mangla, A. T., Nes, W. D., and Waterman, M. R. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96**, 8937–8942
- Podust, L. M., Poulos, T. L., and Waterman, M. R. (2001) *Proc. Natl. Acad. Sci. U. S. A.* **98**, 3068–3073
- Bille, J. (2000) *Int. J. Antimicrob. Agents* **16**, 331–333
- McLean, K. J., Marshall, K. R., Richmond, A., Hunter, I. S., Fowler, K., Kieser, T., Gurucha, S. S., Besra, G. S., and Munro, A. W. (2002) *Microbiology* **148**, 2937–2949

17. Brennan, P. J., and Nikaido, H. (1995) *Annu. Rev. Biochem.* **64**, 29–63
18. Andersen, J. F., Tatsuta, K., Gunji, H., Ishiyama, T., and Hutchinson, C. R. (1993) *Biochemistry* **32**, 1905–1913
19. Lamb, D. C., Fowler, K., Kieser, T., Manning, N., Podust, L. M., Waterman, M. R., Kelly, D. E., and Kelly, S. L. (2002) *Biochem. J.* **364**, 555–562
20. McLean, K. J., Cheesman, M. R., Rivers, S. L., Richmond, A., Leys, D., Chapman, S. K., Reid, G. A., Price, N. C., Kelly, S. M., Clarkson, J., Smith, W. E., and Munro, A. W. (2002) *J. Inorg. Biochem.* **91**, 527–541
21. Mowat, C. G., Leys, D., McLean, K. J., Rivers, S. L., Richmond, A., Munro, A. W., Lombardia, M. O., Alzari, P. M., Reid, G. A., Chapman, S. K., and Walkinshaw, M. D. (2002) *Acta Crystallogr. Sec. D* **58**, 704–705
22. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol.* **276**, 307–326
23. Knight, S. D. (2000) *Acta Crystallogr. Sec. D* **52**, 42–47
24. Otwinowski, Z. (1991) in *Isomorphous Replacement and Anomalous Scattering* (Wolf, W., Evans, P. R., and Leslie, A. G. W., eds) Daresbury Study Weekend, nos. DL/SCI/R32, Warrington, UK
25. Perrakis, A., Morris, R. M., and Lamzin, V. S. (1999) *Nat. Struct. Biol.* **6**, 458–463
26. Roussel, A., and Cambillau, C. (1991) in *Silicon Graphics Geometry Partners Directory 86*, Silicon Graphics, Mountain View, CA
27. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) *Acta Crystallogr. Sec. D* **53**, 240–255
28. Cupp-Vickery, J. R., Garcia, C., Hofacre, A., and McGee-Estrada, K. (2001) *J. Mol. Biol.* **311**, 101–110
29. Cupp-Vickery, J. R., Han, O., Hutchinson, C. R., and Poulos T. L. (1996) *Nat. Struct. Biol.* **3**, 632–637
30. Ravichandran, K. G., Boddupalli, S. S., Hasemann, C. A., and Peterson, J. A. (1993) *Science* **261**, 731–736
31. Yano, J. K., Koo, L. S., Schuller, D. J., Li, H. Y., de Montellano, P. R. O., and Poulos, T. L. (2000) *J. Biol. Chem.* **275**, 31086–31092
32. Muskett, F. W., Kelly, G. P., and Whitford, D. (1996) *J. Mol. Biol.* **258**, 172–189
33. Mortuza, G. B., and Whitford, D. (1997) *FEBS Lett.* **412**, 610–614
34. Modi, S., Primrose, W. U., Lian, L.-Y., and Roberts, G. C. K. (1995) *Biochem. J.* **310**, 939–943
35. Ost, T. W. B., Miles, C. S., Munro, A. W., Murdoch, J., Reid, G. A., and Chapman, S. K. (2001) *Biochemistry* **40**, 13421–13429
36. Raag, R., Martinis, S. A., Sligar, S. G., and Poulos, T. L. (1991) *Biochemistry* **30**, 11420–11429
37. Yeom, H., Sligar, S. G., Li, H. Y., Poulos, T. L., and Fulco, A. J. (1995) *Biochemistry* **34**, 14733–14740
38. Noble, M. A., Miles, C. S., Chapman, S. K., Lysek, D. A., Mackay, A. C., Reid, G. A., Hanzlik, R. P., and Munro, A. W. (1999) *Biochem. J.* **339**, 371–379
39. Li, H. Y., and Poulos, T. L. (1997) *Nat. Struct. Biol.* **4**, 140–146
40. Voiechovsky, J., Chu, K., Berendzen, J., Sweet, R. M., and Schlichting, I. (1999) *Biophys. J.* **77**, 2153–2174